DOCUMENT RESUME

ED 271 010 FL 015 775

AUTHOR Abdrabou, Abdelrahman A.
TITLE Achievement Testing and Foreign Language Teaching.
PUB DATE Apr 84
NOTE 20p.; In: Proceedings of the Annual Linguistics Conference (3rd, Irbid, Jordan, April 1-3, 1984); see FL 015 768.
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; Aptitude Tests; Criterion Referenced Tests; Diagnostic Tests; *Language Tests; Learning Processes; *Reading Comprehension; Second Language Learning; *Test Use

ABSTRACT

It is argued that the close relationship between learning and testing is best exemplified in achievement tests, which are designed both to reinforce learning and to provide feedback about the effectiveness of teaching. Of the four types of achievement tests (placement, formative, diagnostic, and summative), two (formative and diagnostic) are examined in greater depth, and their rationales, functions, and the principles underlying their construction are analyzed. The instructional value of achievement tests and their relationship to motivation, retention, and feedback are reviewed. Reading comprehension testing is discussed in this context, and a sample test designed to enhance learning as well as test skill is presented. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ACHIEVEMENT TESTING AND FOREIGN LANGUAGE TEACHING

Abdelrahman A. Abdrabou
Sana'a University

ED271010

An Overview:

Foreign language tests can be classified into two major categories depending upon the purpose for which they are used. The first category includes achievement and diagnostic tests where the purpose is to assess "achieved" or current performance. The second category encompasses proficiency and aptitude tests. Rather than assessing present performance, proficiency and aptitude tests are basically designed to predict future performance. Within the above limitations and defined purposes, it is reasonable to agree with Spolsky's definition (1968:88) of achievement and diagnostic tests as an essential procedure for the "classroom teacher who wishes to find out how effective his teaching and the students' learning have been or to discover what needs to be taught." In short, the first class of tests (the achievement and the diagnostic) can be utilized for devising appropriate syllabuses to improve present performance, the second class (the proficiency and the aptitude) is used for making an educated guess about a person's chances of academic success in a foreign language.

For the purpose of structuring this paper, it seems extremely important that we establish working definitions for commonly used terms such as "achievement" and "proficiency." Briere (1971, p. 385) defines "achievement" in language performance as the "extent to which an individual student has mastered the specific skills or body of information which have been presented in a formal classroom situation." The problem with the foregoing definition is that it limits instruction to formal classroom situations and it renders "achievement" to mastery of "specific skills." This is a very narrow view of language classes which in principle should be informal. It is also a very narrow view of achievement in language performance which should go beyond

FL015775

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Abdelrahman Abdrabou-Sallam

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

the simple level of recalling "a body of information" to the level of analysis, interpretation, synthesis and eventually application. This might seem to be a very ambitious goal to achieve within the constraints of achievement tests. It is ambitious, true, but it is feasible. It is feasible if we as teachers assume the responsibility for transferring such complex skills to our students during instruction.

On the other hand, Briere (1968, p. 386) defined proficiency as "the degree of competence or the capability in a given language demonstrated by an individual at a given point in time independent of a specific textbook, chapter in the book, or pedagogical methods." The key word in the foregoing definition is the word "independent." Whereas achievement tests are content-bound, proficiency-tests are content-free. This does not mean that proficiency tests do not have a content. It simply means that the sampling items of a proficiency test are to be based on an overall f-ctor analysis of English language skills rather than on a specific content or specified skills.

Even though proficiency tests are very important for assessing the terminal behavior of foreign language learners, this study will, however, limit itself to the issue of achievement testing. Our interest in achievement testing springs from our unfailing realisation that there is a one-to-one relationship between teaching and testing. This relationship can best be exemplified by means of achievement tests. This is simply because achievement tests are primarily constructed to work in two different directions: to reinforce learning, and to provide feedback about the effectiveness of teaching. This argument will be carried a little bit further in the third quarter of this paper. The following discussion will focus on the different types of achievement tests: their rationale, function, and the principles underlying their construction. Of all four types of achievement tests, a special consideration will be given to mastery and diagnostic tests. Our discussion of mastery and diagnostic tests will draw heavily on their possible use for reinforcing reading

comprehension skills and for facilitating difficulties pertaining to reading comprehension.

Language Achievement Tests: Rationale, Function and Principles

Most achievement tests are prepared, administered, scored and interpreted by classroom teachers. They are, therefore, called classroom tests. Their major function is to measure student achievement and thus to contribute to the evaluation of his learning progress and attainment. This is a matter of considerable importance. To argue, as some critics of testing have argued, that what a student knows and can do is more important than his score on a test or his grade in a course implies quite incorrectly in most cases that the two are independent or unrelated. This is absolutely erroneous, because test scores are meant to measure reward and reinforce effective study, penalize unproductive effort, and/or discourage lack of effort. Tests can, and often do, help teachers to give more valid, reliable grades in connection with students' educational achievement.

Test constructors often face two major problems. The first is to determine what to measure. The second is to decide how to measure it. How well the first problem is solved largely determines what is sometimes called the relevance of the test. Robert L. Ebel (1965, p. 22) defines relevance as: "The apparent or obvious logical relationship between what the process of testing requires the student to do and what the process of education undertook to teach him to do." A test is relevant if it strives to strike a balance between what has been taught and what is expected to be learned. The gap between the two can be approximated by appropriate selection of test items. Within the context of foreign language testing, we should first determine the basic areas to be sampled. What are the language features and skills to be tested? How relevant to content are they? And how meaningful to the learner can they be? The problem of relevancy can be solved by clearly identifying goals in a carefully designed language matrix. Test items can be extremely meaningful

if the teacher knows what his students will need the language for. Exploration of the needs of the learner is crucial to the issue of meaningfulness of test items. An EFL student might want to study English to do external readings on subjects such as medicine or engineering of which there is not enough literature in his native language. For such a type of student the teaching of English should focus on sharpening reading comprehension skills, and achievement tests should aim at assessing the degree of mastery of such skills. It is out of proportion and certainly meaningless to construct tests for such type of learners that aim at measuring oral or aural skills.

The extent to which achievement tests contribute to their intended goals is determined largely by the principles underlying their construction and use. The following discussion of achievement test principles is drawn heavily from Norman E. Gronlund's book, Constructing Achievement Tests (1977, pp. 7-14).

> (1) Achievement tests should measure learning abilities that are in harmony with the goals of instruction. Therefore what comes first in constructing an achievement test is not the construction of test items, but rather the identification of the learning abilities to be measured. This is the most crucial stage in the construction of achievement tests because much of what we understand as "learning abilities" is intangible and does not fit the rigid definitions of what are so called "behavioral objectives." The foregoing argument is not meant to underestimate the value of the language testing matrix, because that is a step toward approximating any discrepancy between teaching and testing. A language testing matrix should be utilized as a guideline for the selection of test items, but not as a constraint for all learning goals.

BEST COPY AVAILABLE

(2) Achievement tests should measure a representative sample of the content and subject matter that have been explicitly or implicitly taught. This allows for more of a wide range of goals to be assessed. This is important because if we cannot ensure that an adequate sample of content and subject matter is incorporated in our test, we cannot tell with certainty whether the test result is a reliable measure of achievement or not. In achievement tests, adequate sampling is a matter of greatest concern. For example, is a spelling test of twenty or thirty words a good example of the learner's spelling ability? Raising such kinds of questions is very important for the understanding of the strength and limitations of the whole process of evaluat on.

(3) Achievement tests should include the types of test items that are most appropriate for measuring the intended goals. The major classifications of achievement test items are as follows:

A. Supply-type (student supplies the answer).
   1. Essay - extended response, e.g., free composition.
   2. Essay - restricted response, e.g., controlled composition.
   3. Short answer (word or phrase).
   4. Completion, e.g., fill in blanks or cloze-tests.

B. Selection-type (student selects the answer).
   1. True-false.
   2. Matching.
   3. Multiple choice.

Selecting the appropriate type of test items depends mainly upon the nature of the course and the purpose of the test. In general the supply-type is more appropriate for measuring productive skills such as writing and/or speaking, whereas the selection type is the best fit for measuring receptive skills such as reading and listening comprehension. In the closing section of this paper, we will see how multiple choice tests can be utilized for measuring reading comprehension.

(4) The results of achievement tests should be interpreted only within the specific context and purpose of the test. For instance if an achievement test is specifically designed to assess prerequisite skills at the beginning of instruction (placement test), it will be of no great value to pinpoint difficulties during instruction or to measure goals at the end of instruction. This is simply because placement-tests are - by definition - intended to measure the minimum essentials of a prerequisite unit of instruction or limited set of required skills. Placement-tests tend to be very broad and relatively easy and if used for other purposes, their results might not be valid. The same holds true for diagnostic and summative tests.

Types of Achievement Tests:

Achievement tests can be classified into four basic types of classroom evaluation. These types are: (1) placement evaluation, (2) formative evaluation, (3) diagnostic evaluation, and (4) summative evaluation. Since most achievement tests are teacher-made tests, this classification provides a convenient frame of reference for teachers to consult in planning classroom

tests. Even though the four types of tests are extremely important, we will, however, limit ourselves to formative and diagnostic testing. The reason for this is that we are interested in the interrelationship between learning, teaching, and testing which can best be illustrated by formative and diagnostic evaluation.

Formative Testing:

Formative tests are given periodically during instruction to reassure teachers that what has been taught is understood. A formative test typically covers some predefined segments of instruction. For example in a grammar class, we may want to make sure that our students are capable of using and differentiating between the simple present and the simple past. In such a case, a formative test focusing primarily on these language features is in order. Due to their heavy dependence upon predefined content, formative tests are sometimes referred to as criterion-referenced mastery tests. This is in contrast with proficiency tests which are usually norm-referenced. Ideally formative tests should be constructed in such a way that corrective measures are considered to ensure optimal mastery.

Diagnostic Testing:

Diagnosis of persistent learning problems is an indispensable task for any teacher to undertake. It is indispensable because we always strive for optimal learning. The diagnostic test takes up where the formative test leaves off. If the learner does not positively respond to the corrective feedback of formative testing, a more detailed search for the source of learning errors is called for. Diagnostic tests provide invaluable insights into the nature of learning difficulties since they are constructed in light of the most sources of errors encountered by learners. A language teacher should, therefore, be able to use the information provided by the diagnostic test to determine why the student is deficient in a particular area

of foreign language study. For example, there may be cases in which students cannot use the auxiliary verb 'do' properly in interrogative and negative sentences, or cannot handle the continuous present form in their speech or writing. The search for the sources of errors might entail a study of the learner's native language. In such a case, the literature on contrastive linguistics and contrastive rhetoric might be of great help for the teacher to unveil the underlying causes of the learning difficulties. Error analysis of students' performance on a diagnostic test should ideally encourage the teacher to devise more effective methods of teaching to offset learning problems.

Though we mentioned earlier that diagnostic tests are by definition and by function classroom tests, yet there are plenty of standardized diagnostic tests dealing with persistent learning problems pertaining to specific foreign language communities (e.g., Spanish-speaking community).

The Instructional Value of Achievement Tests:

Achievement testing plays a prominent role in all types of instructional programs. Like any other form of evaluation, the emphasis in achievement testing is on measuring learning, but it should not be implied that this is the ultimate goal of such tests. As with teaching, the main purpose of achievement testing is not only to evaluate, but also to improve learning. In order to realize the full potential of achievement tests as learning aids, it is necessary to make testing an integral part of the instructional process. Testing should be considered before and during the various stages of teaching. From the beginning of instruction to the end, there are numerous decisions that teachers must make. Testing can improve the effectiveness of many of these decisions by providing more objective information on which to base the judgments. At the beginning of instruction, teachers often want to know to what extent the students possess the prerequisite skills needed to start instruction. A well designed placement test can provide

teachers with more reliable information concerning learning readiness and save them the trouble of making erroneous assumptions. Under these conditions, placement tests provide an invaluable aid for placing each student at the most appropriate level of instruction.

During instruction, the main concern of teachers is with the learning progress being made by students. Their concern can best be satisfied by means of formative tests which can be given at the end of each unit of instruction. The effect of such tests on learning cannot be overlooked. This is simply because the ultimate goal of formative testing is to improve learning. If the majority of the students fail a test item or set of items, the material ought to be retaught for the whole group. When a minority of students experience learning failures, a much more individualized or programmed instruction is sought to facilitate learning.

When a student's learning problems become so persistent that they cannot be resolved by the corrective measures of formative testing a more intensive study of the students' learning difficulties is called for. It is at this point, as we mentioned earlier, that the diagnostic test is most useful. The formative test determines whether a student is in good command of the subject matter being taught, and if not, prescribes how to deal with the learning failure. The diagnostic test is designed to probe deeper into the causes of learning deficiencies that are left unresolved by formative testing. All we are attempting to do here is to show that formative and diagnostic tests can contribute to improve student learning during instruction. And by so doing, formative and diagnostic tests help teachers to make intelligent and wise decisions about the effectiveness of their teaching methods and the appropriateness of their subject-matters.

In addition to improving instructional decisions, the use of tests, in general and achievement tests in particular, can aid learning by:

(1) Improving student motivation;

(2) Increasing retention and transfer of learning; and

(3) Providing feedback concerning instructional effectiveness[1].

Each of these will be discussed in turn.

Testing and Motivation:

Periodic testing motivates students by providing them with specific goals to work for, and by providing them with immediate feedback concerning their achievement and learning progress. If tests are carefully constructed and if goals and expectations are clarified and shared between student and teacher, there is no reason to worry about the side effects of testing such as anxiety and fear. Fear and anxiety can be reduced or even utilized to generate greater learning activity. Thus, properly constructed tests can motivate students to work toward the specific goals of a course by inviting greater learning activity, by directing this activity toward the intended learning objectives, and by providing prompt knowledge of results.

Periodic testing can also aid student motivation by helping individuals understand themselves better so that they can make more intelligent decisions about their future. Formative testing, for instance, and feedback of the results can help students gain insight into the things they can do well, the misconceptions that need correction, the degree of skills they have in various areas, and the like. Such information provides the student with a more objective basis for planning a study program, for selecting future educational experiences, and for developing self-evaluation skills. With better self-understanding, the degree of motivation among students is likely to increase.

Testing and Retention:

Since tests tend to direct students' learning efforts toward specific goals, they can be used as tools for increasing the retention and transfer of classroom learning. In general if tests are oriented toward the evaluation of learning at the

BEST COPY AVAILABLE

understanding, application, and interpretation levels, the retention of skills and subject-matters are likely to remain longer and to have greater transfer value than knowledge at the recalling level. Thus tests, in this sense, can be used to supplement and complement our teaching efforts in these areas and thereby increase the likelihood that learning will be of greater permanent value to students. To achieve this goal, test-constructors are often urged to de-emphasize recognition or simple recall of isolated factual details. They are advised to shun tests of this type in favor of tests that emphasize comprehension, interpretation, application, analysis, synthesis, and evaluation. They are cautioned to guard against items which can be answered on the basis of sheer rote memory because this can be counterproductive to long-term retention, and to the transfer of learning. In the last section of this paper, we will see how these principles can be incorporated in reading comprehension tests.

Testing and Feedback:

We have already explained how the test results can be used to improve and enhance learning. But this is not the whole thing. Information provided by test results can be utilized to evaluate various aspects of teaching. It can help to determine the extent to which the content of the curriculum is realistic, whether methods and teaching strategies are appropriate, and what course of action can be taken in response to such questions. When the majority of the students do poorly on a test - particularly a formative test - this may partly be the fault of the students, but the difficulty is more apt to be found in the instruction. The teacher may be striving for goals that are, for some reason, unattainable by the students, or he may be using ineffective methods to communicate content to his students. With this broad view of testing, it is reasonable to argue that teaching without testing is unthinkable, and testing is just as important to learning as teaching is.

The following discussion will demonstrate how the principles and objectives of achievement testing can be manifested in reading comprehension tests.

Testing Reading Comprehension:

Most foreign students study a language with the end goal of reading and understanding the printed page. This emphasis on reading comprehension is quite justifiable particularly if we bear in mind that most foreign students study English to be able to read the advanced literature in their field of specialization which might be unobtainable in their native language. In addition to this, overemphasis on speaking and pronunciation might be a waste of time as most of the foreign students might not have the opportunity to travel abroad to practice their speaking skills.

Reading comprehension tests often include as one of their parts a vocabulary test. The reason for this is that satisfactory comprehension of any printed materials depends in part on knowledge of vocabulary. For effective reading comprehension, lexical mastery is, therefore, indispensable. However, there are other skills besides lexical mastery upon which understanding and reading comprehension depends. Skills such as the ability to make inferences, to analyse, to interpret, and to pass on a value judgment are indispensable to effective reading. The question is: Are these skills testable? The answer is positively yes. It is my contention, from the start of this paper, that anything that is teachable is also testable and vice versa.

Passages in reading comprehension tests should be meaningful and relevant to the examinees. Insofar as we are dealing with achievement testing, reading passages should provide some contextual clues to meaning to help the learners see the point. These characteristics and others will be demonstrated in the following sample of a reading comprehension test which is specially designed for advanced EFL learners. The author assumes that passages similar to the ones presented in the test were

previously studied in class and that most of the vocabulary items are within the capacity level of the students.

Reading Comprehension Sample Test

Time: 10

Directions:

Below the following passage, you will find questions or incomplete statements about the passage. Each statement or question is followed by a number of words or expressions. Select the word or expression that best fits each statement or answers each question in accordance with the meaning of the passage. Students should manage to read the passage as quickly as possible.

Passage:

The teaching of a foreign language, unlike that of content subjects, can be, and frequently is, extremely frustrating and wasteful. The most immediate example of this is the so-called Intensive English Language Program at Sanaa University. Students enter the university after having done about six years of English at the intermediate and secondary level. Yet a large number of them are not in a position to produce a simple sentence in English. What is even more disappointing and discouraging is that there is almost no improvement in the competence, the ability to produce and communicate in English, of a sizable number of these students even after four years of being taught English at the University.

It is customary to blame this state of affairs on the teachers, the courses, and the methods used for teaching English as a foreign language. The teachers, in their turn, rationalise their failure and frustration by alluding and referring indirectly to the physical inadequacies of their teaching situation, e.g., inordinately large classes, unsuitable rooms and furniture, insufficient hours of teaching, and ineffective administrative control over students' class attendance and performance.

Example:

The passage is principally concerned with:

(A) Teaching of English at Sanaa University.

(B) The inadequate facilities of Sanaa University.

(C) The consequences of ineffective control over students' attendance.

(D) Problems of teaching English as a foreign language at Sanaa University.

Choice (D) is correct and most specific.

The Questions:

(1) Which of the following can be inferred about Sanaa University?

A) It only admits students who had six years of English language instruction.

B) It offers intensive English language courses for four years.

C) It does not have enough rooms and furniture.

D) It needs more effective and stimulating methods for organizing and directing the English as a Foreign Language Program.

(2) According to the author, Sanaa University students had instruction in English before joining college for:

A) One year.

B) Two years.

C) Six years.

D) Had no instruction in English at all.

(3) By "this state of affairs," the author refers to:

A) Ineffectiveness of the administration.

B) Teachers' recklessness.

C) Insufficient hours of teaching.

D) Poor performance on the part of the students.

(4) The word rationalize in the passage is used to mean:

A) Explain in a logical way.

B) Verify.

C) Illuminate.

D) Refer.

(5) By stating that the teaching of a foreign language is "unlike that of content subjects," the author means that:

A) The English language is different from other content subjects.

B) The teaching of English differs from that of other content subjects.

C) The English language lacks content.

D) English is boring and frustrating.

(6) The teaching of English at Sanaa University is retarded because of:

A) Lack of interest on part of the students.

B) Lack of motivation on part of teacher.

C) Lack of cooperation and understanding between teachers, students, and administrators.

D) None of the above.

(7) The author's attitude toward the teaching of English at Sanaa University is:

A) Pessi istic.

B) Optimistic.

C) Neutral.

D) Characterized by a great deal of concern and worries.

(8) By 'alluding' the author means:

A) Indicating.

B) Explaining.

C) Implying.

D) Describing.

(9) The author implies that the teaching of English can be "frustrating" and "wasteful" if:

A) The number of students in class is too small.

B) It is not accompanied by improvement in students' performance and competence.

C) There is no cooperation between teachers and administrators.

D) Both (B) and (C)

(10) The word that can best stand for "inadequacies" is:

A) Deficiencies.
B) Insufficiencies.
C) Inefficiencies.
D) Dissatisfaction.

Test Key:

(1)	D	(6)	C
(2)	C	(7)	D
(3)	D	(8)	C
(4)	A	(9)	D
(5)	B	(10)	A

Test Rationale:

Testing reading for the advanced level does not limit itself to the level of rote memorization; rather it includes skills and techniques for the understanding and analysing of the material at hand. This requires a certain speed-rate and specific reading skills to get the needed information efficiently and rapidly from the assigned reading passage(s).

Our objective in constructing this test is to measure the speed and ease with which advanced EFL learners can encode and decode messages. The emphasis then is not simply on determining the student's control of the basic signalling devices, but also on his understanding of the semantic and syntactic implications of such signalling devices. Speed is unquestionably an important aspect of language mastery.

To achieve this goal, our test includes different types of questions which vary in degree of difficulty. The test questions fall in the following categories.

(1) Questions that require knowledge of specific facts, e.g., Question No. 2.

(2) Questions that require inferences or implications from the reading passage and which are not directly stated in the selection as in Questions No. 1, 3, and 7.

(3) Questions that require synthesis as in No. 6.

(4) Questions that require syntagmatic understanding - that is, questions that require an understanding of the fusion of two or more linguistic signs or elements in a word, phrase, or idiomatic expressions as in Question No. 5.

(5) Questions that require lexical mastery as in Nos. 4 and 8.

By virtue of being formative, this test provides certain contextual clues to help the examinees to understand vocabulary meaning. These contextual clues are presented by means of:

I. Synonyms: as in the sentence, "What is even more disappointing and discouraging is that .....".

II. Definitions: The word is explained or defined by the author as in the sentence "by alluding and referring indirectly .....".

III. Relevancy: The meaning is clarified by the relevancy of the subject matter to the experience of the learner. So far as this test is constructed for EFL students at Sanaa University, its content is related to their own personal experiences. Meaningful content ensures motivation and enhances successful learning.

IV. Summary: An idea or situation expressed in different ways is summed up in one word or expression. The whole phrase, "inordinately large classes, unsuitable rooms and furniture, insufficient hours of teaching, and ineffective administrative control over students' class attendance and performance," is summed up by the expression of "physical inadequacies of their teaching situation."

The idea behind giving such contextual clues is to increase the learner's reading-speed and to improve his/her other reading skills. Clues are provided to help the learner to see the point. It is from this perspective that formative tests can be perceived as an aid to learning. This is, of course, besides the follow-up activity which is based upon the test results, and which is primarily conducted to offset learning difficulties.

Concluding Remarks:

In concluding this paper, I would like to reiterate my proposition that testing is not incompatible with learning, and that if care is taken in test-construction, testing can in fact improve the quality of both teaching and learning. This paper attempts to explain the interrelationship between teaching and learning. A special emphasis is given to achievement tests particularly to formative and diagnostic testing. The last section of the paper focused on reading comprehension tests, their values and characteristics. A sample reading comprehension test was presented at the end of the paper to illustrate how formative reading comprehension tests can be constructed in such a way to aid learning.

BIBLIOGRAPHY

Briere, Eugene J. 1971. "Are We Really Measuring Proficiency With Our Foreign Language Tests?", Foreign Language Annals 4, 385-386.

Bormuth, John R. 1970. On the Theory of Achievement Test Items. The University of Chicago Press.

Carrol, John B. 1961. Research on Teaching Foreign Languages. Univ. of Michigan, Ann Arbor.

Davies, Allan. 1977. "The Construction of Language Tests." Testing and Experimental Methods. J. P. B. Allen and Allan Davies, eds. Oxford University Press, 38-104.

Ebel, Robert L. 1965. Measuring Educational Achievement. Prentice-Hall, Inc.

Fitch, Mildred L. 1951. "Frequent Testing as Motivation Factor in Large Lecture Classes," Journal of Educational Psychology 42, 1-20.

Gronlund, Norman E. 1977. Constructing Achievement Tests. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Lado, Robert. 1964. Language Testing: The Construction and Use of Foreign Language Tests. McGraw-Hill Book Company, New York.

Satah, Shiroh. 1968. The Development and Interpretation of Several Types of Tests With Emphasis on Their Application in Teaching The English Language to Non-Native Populations. An unpublished Masters thesis, University of Kansas.

Spolsky, Bernard. 1968. "Language Testing: The Problem of Validation," TESOL Quarterly 2, 88-94.